Silicon Graphics, Inc.

Heterogeneity in the Grid

Monterey Grid Computing Conference

August 11, 2004

Presented by:

Walter Stewart, Business Development, Canada

Global Co-ordinator Grid Strategy

Heterogeneity in the Grid

FIRST A COMMERCIAL

www.ggf.org

Global Grid Forum 12, Brussels, September 20-23

Heterogeneity in the Grid

THE GRID WILL BE HETEROGENEOUS

Heterogeneity in the Grid

BUT FIRST I’D LIKE TO START WITH A STORY

Heterogeneity in the Grid

.S. United States

Heterogeneity in the Grid

- The First Cars Looked Like Buggies
- The First Aeroplanes Looked Like Steamers and Pullmans

WE HAVE TO BE CAREFUL THAT WE LIMIT THE CONSTRAINTS ON THE FUTURE BY THE WAY THE PAST CURTAILS OUR IMAGINATION

Heterogeneity in the Grid

WE MUST BE ESPECIALLY CAREFUL OF IDEOLOGY

One generation's revolutionaries are frequently the next generation's reactionaries

Heterogeneity in the Grid

- As we replace the infrastructure of 1999, do we uncritically clone it with the 2004/5 version – faster, smaller, cheaper?

- Or do we recognize that there is cause to develop whole new orders and types of infrastructure to take us towards the end of the first decade of the new millennium?

Heterogeneity in the Grid

That cause is data and the fact that we have begun a knowledge economy.

Heterogeneity in the Grid

SO, THE GRID WILL BE HETEROGENEOUS

Heterogeneity in the Grid

BUT LET’S BE CLEAR:

HETEROGENEITY REFERS TO A WHOLE
MORE THAN CONFIRGURATIONS AND
OPERATING SYSTEMS

Heterogeneity in the Grid

- **THE GRID IS THE INFRASTRUCTURE FOR THE KNOWLEDGE AGE**
- **JUST AS WE BUILT INFRASTRUCTURE FOR THE INDUSTRIAL AGE TO MOVE MATERIALS**
- **SO WE ARE BUILDING INFRASTRUCTURE TO MOVE DATA FOR THE KNOWLEDGE AGE**

Heterogeneity in the Grid

- YES THE GRID WILL ENCOMPASS MULTIPLE OPERATING SYSTEMS AND MANY DIFFERENT KINDS OF COMPUTE CONFIGURATIONS

- BUT THAT MIGHT BE DESCRIBED AS MICRO-HETEROGENEITY

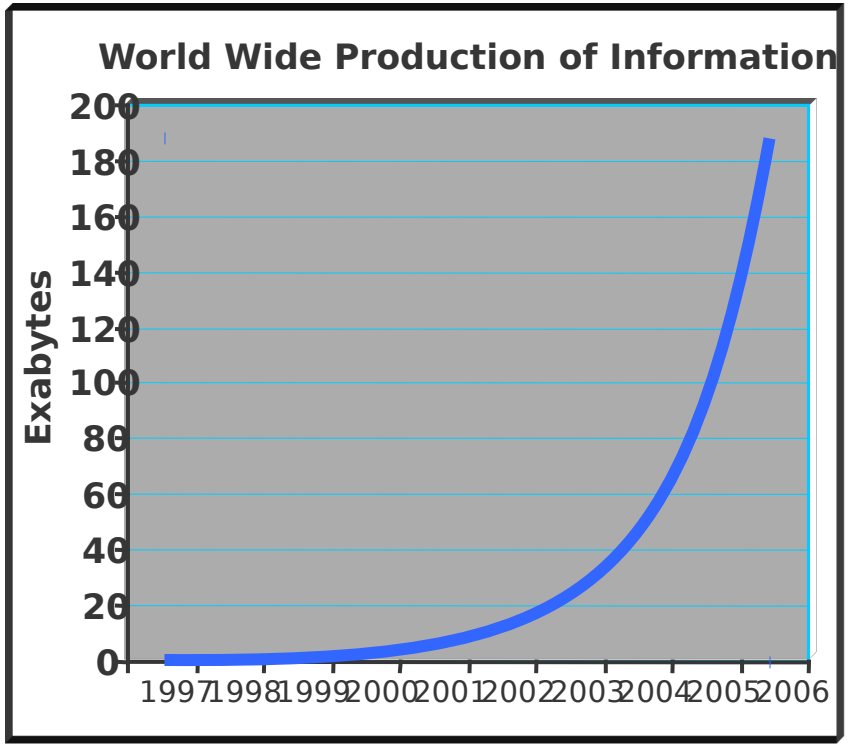
Heterogeneity in the Grid

- MACRO-HETEROGENEITY MEANS THAT THE GRID WILL INVOLVE:
  - Data capture systems
  - Data ingest systems
  - Data management systems
  - Data processing systems
  - Data distribution systems
  - Data visualization systems

Heterogeneity in the Grid

The Data Explosion has become a near cliché but do we understand it?

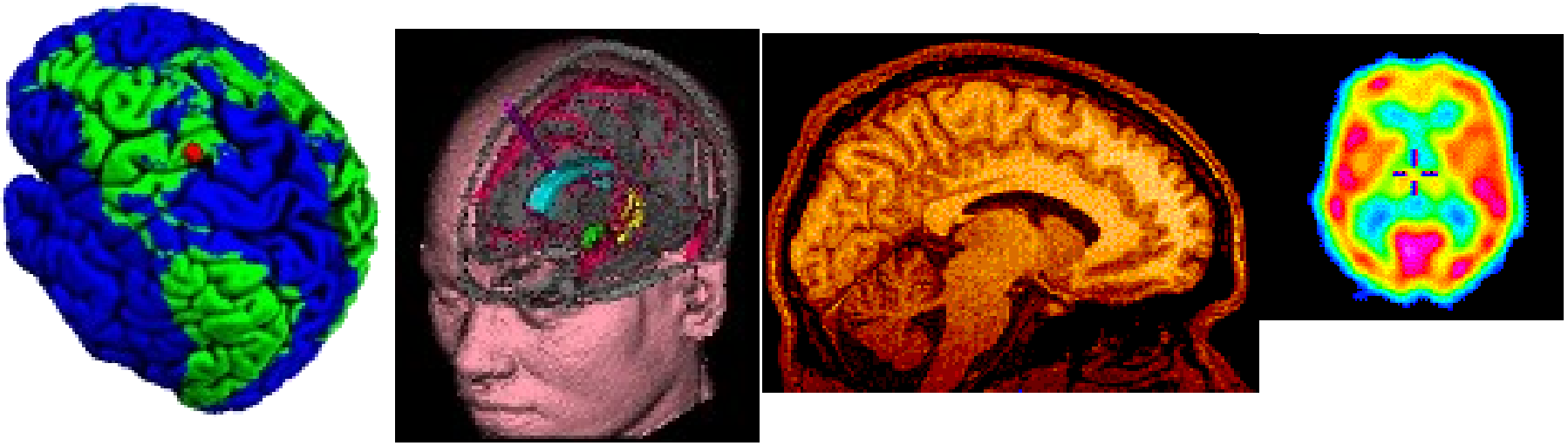
- Energy : 38 TB of seismic data in a single file system, multi-terabyte mirrors over a wide area network
- Media: 200+ TB of data, 20 million files
- Weather: 2TB of data ingested every day
- Customers with Petabytes of data that need to be managed
- Multi-terabyte NAS farms are increasingly common



Source: Gartner Group, June 2003

Heterogeneity in the Grid

Montreal Neurological Institute



Images courtesy of the Montréal Neurological Institute

800 Gigabyte Brain Model

Heterogeneity in the Grid

FOR FIVE CENTURIES SCIENCE HAS BEEN CONDUCTED IN A REDUCTIONIST MODEL.

WE LIVE AT THE BEGINNING OF AN AGE OF INTEGRATIVE SCIENCE.

Heterogeneity in the Grid

BY ITS VERY DEFINITION, INTEGRATIVE SCIENCE IS BIG DATA SCIENCE

Heterogeneity in the Grid

LINKED, THE NEW SCIENCE OF NETWORKS -

Albert-Laszlo Barabasi,
Emile T. Hoffman Professor of Physics, University of Notre Dame

Heterogeneity in the Grid

RIDING REDUCTIONISM WE RUN INTO THE HARD WALL OF COMPLEXITY

Heterogeneity in the Grid

o world of spring and autumn, birth and dying
The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of the Word.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to GOD.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?
The cycles of Heaven in twenty centuries
Bring us farther from GOD and nearer to the Dust.

T.S. Eliott - Opening Chorus from the Rock

Heterogeneity in the Grid

Both Eliott and Barabasi are saying that we may well have mastery of the pieces, but we remain ignorant of the whole.

Heterogeneity in the Grid

COMPLEXITY NEEDS A DIFFERENT APPROACH AND A DIFFERENT INFRASTRUCTURE

THE STUDY OF COMPLEXITY NEEDS LOTS OF MEMORY AND LOTS OF THAT MEMORY MUST BE SHARED

Heterogeneity in the Grid

Big Data Means Science Needs a New Kind of Infrastructure

Heterogeneity in the Grid

That infrastructure is infrastructure for data. It is, in fact, the Grid.

- **Data Capture – Sensor Networks, Scanners, etc.**
- **Data Ingestion – Storage Area Networks**
- **Data Management – CXFS, DMF, Storage Products**
- **Data Processing – Altix 350, Altix 3000, Origin**
- **Data Visualization – Gemini, Onyx**
- **Data Collaboration – Viz Server, CXFS, and CXFS over WAN**
- **Middle Ware – Globus, Platform, Unicore,**
- **Applications – software suitably licensed**
- **and the networks to connect it all**

sgi

Heterogeneity in the Grid

At SGI, we work to enable access to the power we create in compute, data management, and visualization to diverse O.S. and diverse devices.

Heterogeneity in the Grid

Transparent, Universal,

Access

That's what SGI is working to pr

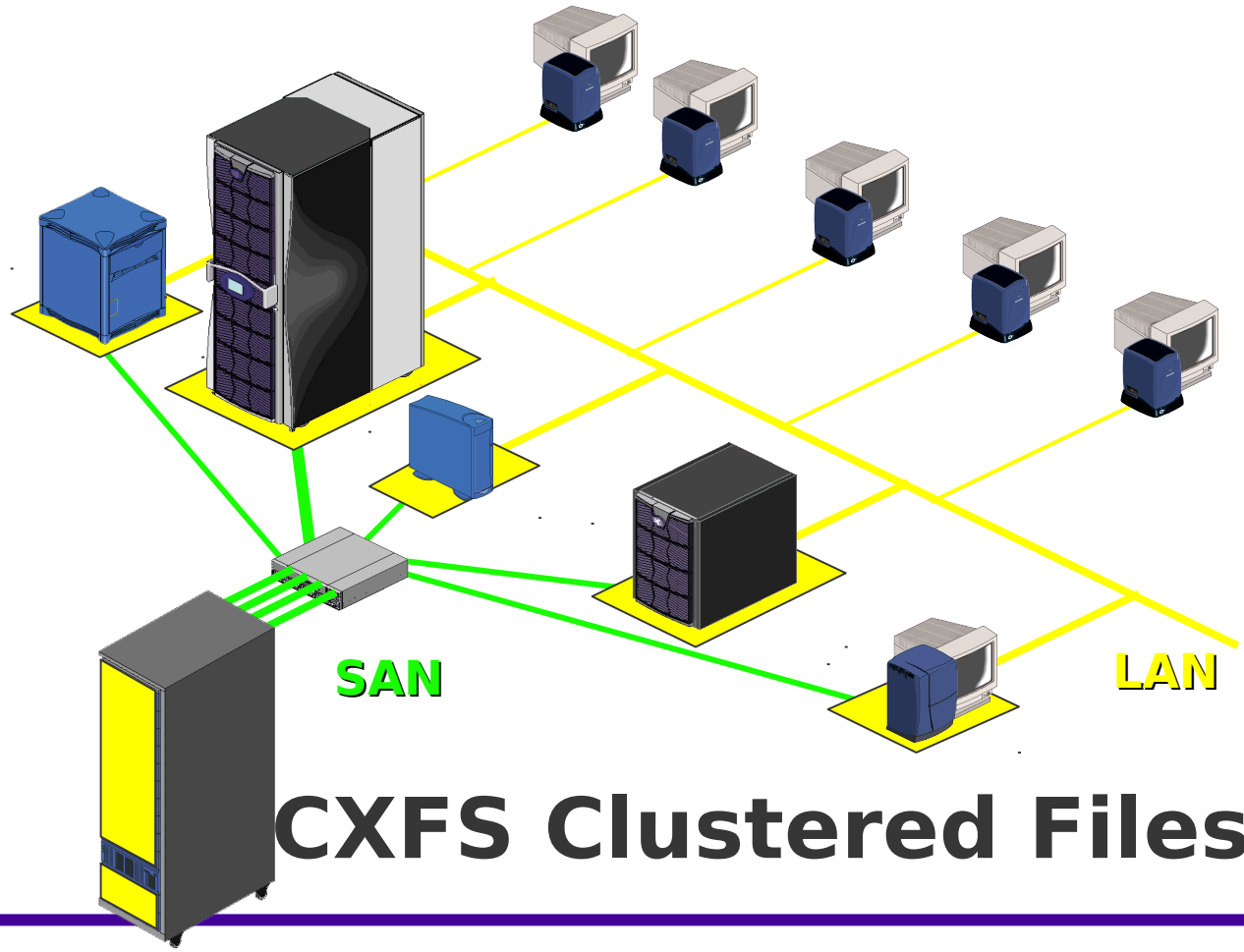
Heterogeneity in the Grid

- SGI customers are actually using:
  - Single filesystem over 75TB
  - Single filesystem > 12GB/sec
  - Single HSM > 3PB
  - Single SAN > 500TB
  - Single NAS > 60TB

Heterogeneity in the Grid

A unique high performances solution:

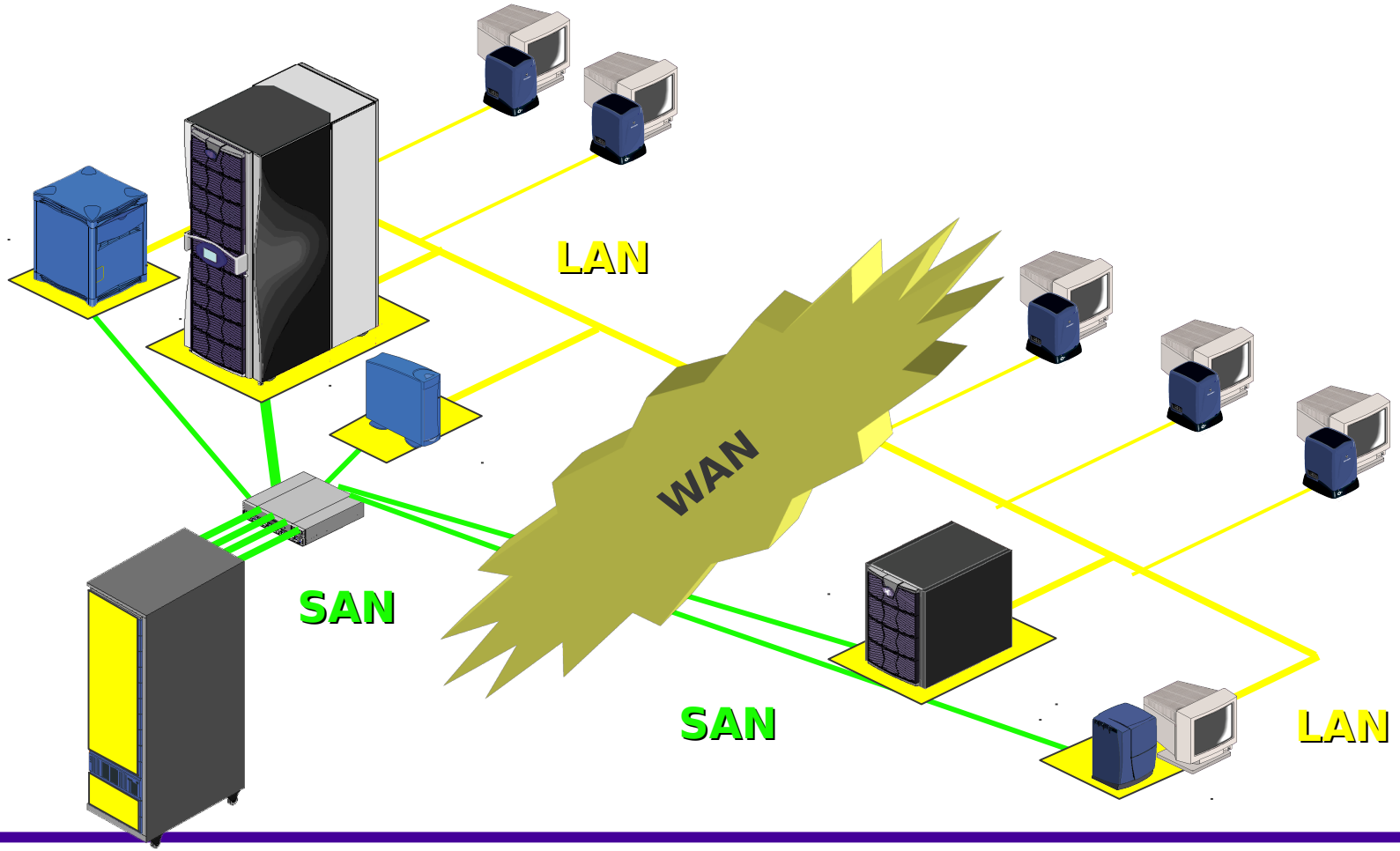
- Each host share one or more volumes consolidated in one or more RAID array.
- Centralized storage management
- High modularity
- True High Performances Data sharing, near local File System performances.
- Fully Resilient (HA)
- Fully POSIX Compliant
- As easy to share files as with NFS, but faster



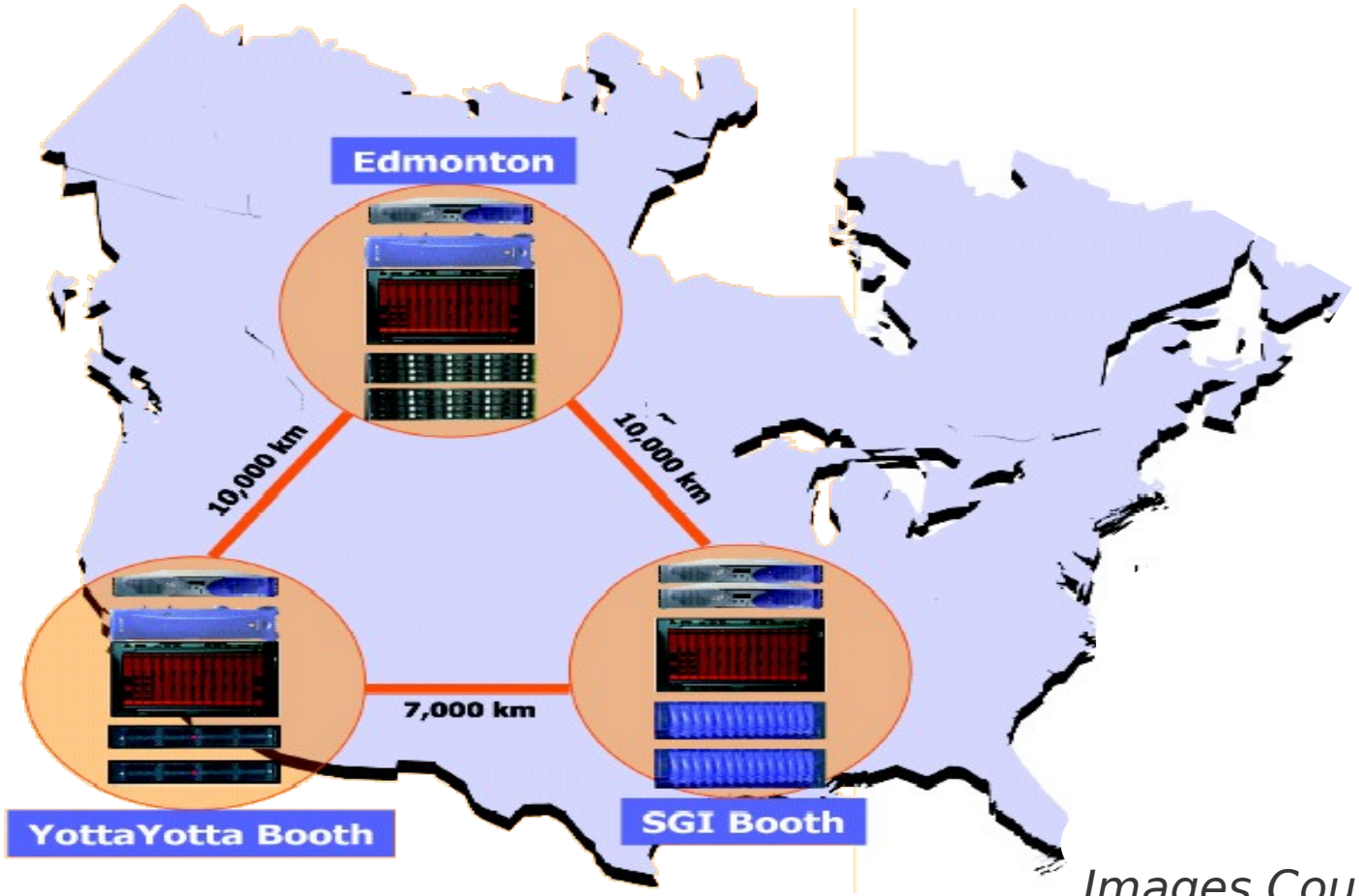
CXFS Clustered Filesyste

sgi

Heterogeneity in the Grid



Heterogeneity in the Grid



Images Courtesy of Yott

Heterogeneity in the Grid

- *CXFS is a high-performance, POSIX compliant, clustered filesystem for Linux, IRIX, Solaris, NT, and AIX*
- *CXFS provides a single filesystem view which spans a cluster and which has the same performance characteristics as a native filesystem (XFS)*
- *CXFS enables multi-system concurrent/sequential file sharing at local file system speeds and with minimal impact to system resources.*
- *CXFS Eliminates the need for NFS/FTP within the shared file system membership*
- *CXFS works over WAN enabling file sharing on the Grid*

Heterogeneity in the Grid

Lessons From History

- the lever, the catapult, the wind, the horse, the steam engine, intern combustion, the jet engine, nuclear power

- All productivity spikes have been a consequence of harnessing unprecedented power

Heterogeneity in the Grid

- We are now engaged in bring the pro from the realm of physical force to th realm of the mind.

- Separating the grunt work from the insight

- We don’t even have commonly understood language to describe what we are about.

Heterogeneity in the Grid

- **Solving the Human IO problem**
  - **it's a question of the medium**
- **Building Communal Intelligence**
  - **it's also a question of the medium**

the oil and gas example

Heterogeneity in the Grid

- **Is the possibility of genuine communal intelligence the basis of the next big productivity spike?**

SGI PROVIDES THE TECHNOLOGY THAT MAKES COMMUNAL INTELLIGENCE POSSIBLE IN HETEROGENEOUS, DISTRIBUTED ENVIRONMENTS

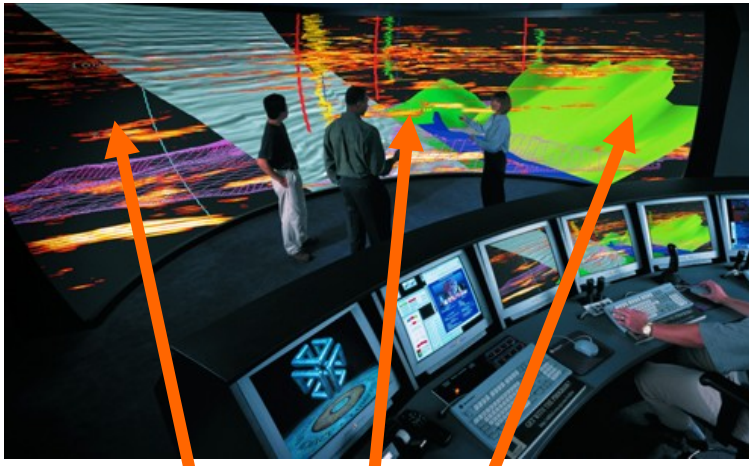
Heterogeneity in the Grid

VISUAL AREA NETWORKING

All collaboration cannot be co-located

Heterogeneity in the Grid

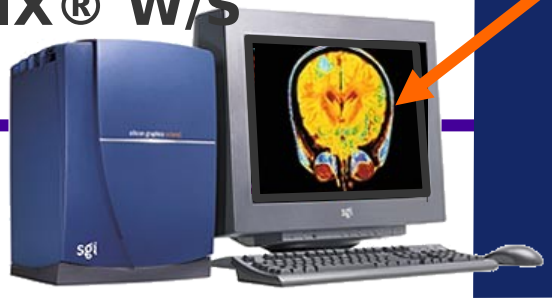
Windows® or
Linux® W/S



Sun W/S

Grid Infrastructure

IRIX® W/S



Heterogeneity in the Grid

You have already heard all sorts of advice in these two days:

- Start small
- Be driven by a business problem
- Avoid “hype”

Heterogeneity in the Grid

I have some advice as well:

- Be clear on your vision
- Root your efforts in a move towards infrastructure
- Avoid the constraints of the present
- Recognize that this is only partly about technology
- Always test any proposed action against how it contributes to building the vision

Heterogeneity in the Grid

- SGI will be honoured to work with you as you build towards that vision.
- We will ensure that that which we bring to the table will co-exist and complement that which you have from other vendors.
- Transparent Universal Access is possible

Heterogeneity in the Grid

Thank You